

日本語を第二言語として学習する成人の 促音認識について

山 根 麻 紀

On adult L2 Japanese learners' perception of moraic consonant Q

Maki YAMANE

1. 序 論

本論では、日本語を第二言語 (second language)ⁱとして学習する成人の音声認識についての実験結果を報告し、ならびに成人第二言語習得一般を音声学的視点から論じる。

成人の言語習得は、子供の言語習得とは以下の(1)と(2)に挙げるような点で異なっている。

- (1) 子供が、通常の言語環境にあって重篤な脳障害等がない場合には、ほとんど個人差なく言語ⁱⁱを習得するのに対し、成人の言語(第二言語)習得には段階的・結果的な個人差がある。
- (2) 成人が習得する言語(第二言語)の最終段階の状態 (final state) は当該言語を第一言語とする話者 (ネイティブ・スピーカー) のものとは異なる場合が多い。特に音声面で両者の違いは顕著で、成人第二言語学習者の発話にはその第一言語 (native language) のアクセントが残る場合が多い。

上記(1)及び(2)の一般的観察から、言語習得には臨界期 (critical period) が存在するという見解がある。臨界期は、ほぼ思春期 (10~12才) 以前に該当し、それ以後の言語習得は、脳の言語習得に関する器官の変化 (または停止) によって困難になるという説である (Lennenberg 1967, Krashen 1973, Curtiss 1977 他)。

このような立場から、成人の第二言語音声習得に関して(3)のような説が主張されている (Eimas 1975, Strange and Dittman 1984 他)。

- (3) 音声認識をつかさどる脳神経系内的器官は、発達の早期段階において、第一言語環境で存在する音声体系のみを識別するよう調整される。臨界期を過ぎて成人になるとその認識システムは固定されてしまい、第一言語に存在しない音声体系を処理するのが非常に困難 (または不可能) になる。

上記(3)の是非はいまだに議論の対象となっている。Miyawaki, Strange, Verbrugge, Jenkins and Fujimura(1975)及びStrange and Dittman(1984)では、日本語を第一言語とする成人英語学習者の /r/ と /l/ の知覚実験が報告されているが、これらの実験結果は(3)の説を支持するものである。すなわち、両研究の被験者は、英語の /r/ と /l/ を識別できなかった。一方で、Pisoni, Lively and Logan (1994) は、実験の前段階にトレーニング・セッションを行った同様の実験で、(3)の説を支持しない結果を報告している。彼らは、トレーニングで認識経験を得ることによって、成人学習者も第一言語に存在しない音声処理できるようになる、すなわち、音声知覚システムは臨界期を過ぎても調整可能である、という結論に達している。

本論で報告される実験は、上に述べたような臨界期以後 (成人) の音声認識システムの可塑性を調査する

のを旨とした。対象としたのは英語を第一言語とする成人日本語学習者で、実験材料は日本語の促音である。促音は、かな書きでは小さい「っ」で表される無音の音単位で、調音音声学的に言えば、無声閉鎖音 (voiceless stop、例えば /p/, /t/, /k/ など) の解放 (release) の保留部分である。促音は、日本語のネイティブ・スピーカーにとっては自明のように処理される音声単位であるが、異なる言語タイミングを有する言語のネイティブ・スピーカーにとってはその限りではない。それは、以下のような理由によると推察される。

日本語では、モーラ (mora) と呼ばれる音声単位が言語タイミングをつかさどっており、それは (4) のように表される。

(4) いっぱいアイスクリームがほしいんです
I (p) pa i a i su ku ri i mu ga ho shi i n de su^u
V Q CV V V V CV CV CV V CV CV CV CV V N CV CV

上記 (4) 中、Cは子音 (consonant)、Vは母音 (vowel)、Qは促音、Nは撥音 (かな書きでは「ん」で表される鼻音) を表している。モーラには、CV、V、Q、Nの4種類があり、全て (4) のようにかなに対応している。日本語のモーラは単語内で言語タイミングをつかさどる。例えば、「ほん」(ho-n, CV N) は2モーラ語、「えほん」(e-ho-n, V CV N) は3モーラ語だが、各々の単語の長さの比はモーラ数に比例してほぼ2:3であることが、過去の研究によって明らかにされている (Port, Dalby and O'Dell 1986, Han 1994, Yamane 1997 他)。ここで留意すべきは、日本語では促音 (Q) や撥音 (N) のように、母音をともしない音声単位も一定の「長さ」を持ち、言語タイミングの統制に参加している、ということである。

一方、英語では (5) のように、モーラではなくストレス (stress) が言語タイミングをつかさどっている。

(5) W S W S W S W
or won to lo to aiskrim
I want a lot of icecream

英語の語は母音を中心に構成されたシラブル (syllable) という単位で成り立っており、上記 (5) ではストレスのあるシラブルの母音上に S (strong)、ストレスのないシラブルに W (weak) があてられている。英語は、ストレスのあるシラブル (S) が、時

間的にはほぼ等間隔で起こるような言語タイミングを有する (Carr 1999 他)。つまり、英語ではストレスが「拍子」のような役割をして言語タイミングを統制しているのである。ここで留意すべきは、ストレスは音の高さ (pitch)、大きさ (loudness)、長さ (duration) の総和であり、これらは有声エネルギーを持った音、すなわち母音の上にしか具現され得ない、ということである。

要約すれば、日本語のタイミング構成要素には、促音のように母音を含まないものもあるが、英語のタイミングを統制するストレスは、特性上母音に現れるので、そのタイミング構成要素は必然的に母音を含むことになる。

ここで成人の第二言語習得の問題に立ち戻り、上で述べた成人日本語学習者の促音処理にともなう困難を鑑みれば、次のような可能性が考えられる。第一言語が英語である場合、音声認識システムはまず英語の音声認識用に調整される。言語タイミング処理の点から言えば、ストレスのある母音で構成されるリズムを認識するよう調整される。彼らが日本語を学習する際には、これとは異なる言語タイミングを習得しなければならない。つまり、英語ではタイミング処理の際に使用していない母音の欠如した音声単位 (促音) を認識し、知覚システムを修正しなければならない。上記 (3) で述べた臨界期論が妥当だとすれば、成人学習者にとってそのような修正は、非常に困難、または不可能であろう。妥当でないとすれば、Pisoni, Lively and Logan (1994) の報告するように、認識経験を積むことで修正は可能であろう。

以上をふまえて、本論では以下の (6) のような仮説を立てて実験を行った。これは便宜上、上記 (3) の臨界期論を妥当と仮定してたてられたものである。

(6) 実験仮説

英語を第一言語とする成人日本語学習者の音声認識システムは、英語の音声体系を識別するようにのみ調整され、固定されている。

すなわち、言語タイミング処理の点では、英語のストレス・タイミングに基づく処理だけが可能である。ストレスは母音上にしか体现されないことから、母音を持たない音声単位である日本語の促音の有無は当被験者にとって処理不能である。したがって、その促音認識パターンは、

日本語のネイティブ・スピーカーのパターンと一致しない。

2. 実験

実験は1999年8月、ミネソタ大学秋田校夏期日本語集中講座の中級クラスで行われた。被験者となった16人はいずれも英語を第一言語としており、日本各地にALT (assistant language teacher) として赴任している。年齢は24才から30才で、全員が来日して1年、来日以前に日本語を学んだ経験はなく、来日してからも正規のクラスで学んでいない。当集中講座のクラス分けテストに採用された日本語能力検定試験3級では、被験者全員が40%以上60%未満の正答率の範囲に入っている。

実験材料の言語音には、「いき」(iki) と「いっき」(iQki)、「いた」(ita) と「いった」(iQta) の、促音の有無を差異とする2つのミニマル・ペアが用いられた。この4つの語各々について、3人の日本語ネイティブ・スピーカーによって通常のテンポで発話されたトークン30個のうち、母音部・子音部・語全体の長さが最も平均値に近いものが1つ選ばれた（以下これを基本トークンと呼ぶ）。それを音声分析ソフト (Signalize 3.1) にかかけ、子音部を編集した。促音を含まない基本トークンには、子音部に10ミリセカンド (msec) ずつ無音部をバーストしていった（以下この過程でできたものを中間トークンと呼ぶ）。この作業を、ミニマル・ペアのもう一方の、促音を含む語の値になるまで繰り返した（以下これをターゲット・トークンと呼ぶ）。例えば、「いき」(iki) を基本トークンとした場合、その子音 /k/ の部分は、母音 /i/ と /a/ に挟まれた空白部分として表れるが、それを10msecずつのばしていったことになる。その結果、「いき」は、厳密

に言えば「いき」と「いっき」のいずれでもない中間トークンをいくつか経て、「いっき」という音として認識されるまで編集される。促音を含む基本トークンからは、それとは逆に子音部を10msecずつカットしていった。この場合、例えば「いっき」の子音部は中間トークンにおいて次第に短くなり、ターゲット・トークン「いき」に至る。ひとつの基本トークンからそれぞれ11の中間トークンと1つのターゲット・トークンが作成された（以下これらをまとめてシリーズと呼ぶ）。このようにして、上の4つの語を各々基本トークンとした4シリーズの言語音刺激が作成されたことになる。すなわち、①「いき」→「いっき」シリーズ、②「いっき」→「いき」シリーズ、③「いた」→「いった」シリーズ、④「いった」→「いた」シリーズ、である。

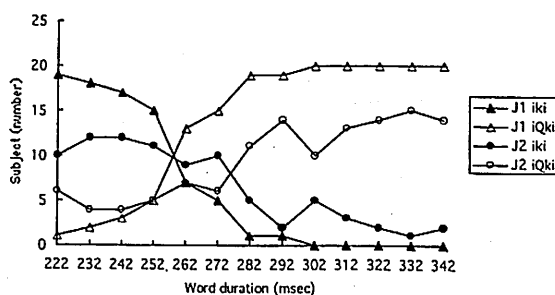
各シリーズ内でトークンの順序を無作為に入れ替え、トークン間の余白を2秒とって、オーディオテープに移し替えたものが被験者にテープレコーダを通して呈示された。呈示は静かな教室内で被験者全員に一斉に行われた。

被験者には、二者選択式の解答用紙が配られ、各々のトークンがミニマル・ペアのどちらに聞こえるか答えるよう指示が出された（以下これをラベリング (labeling) と呼ぶ）。

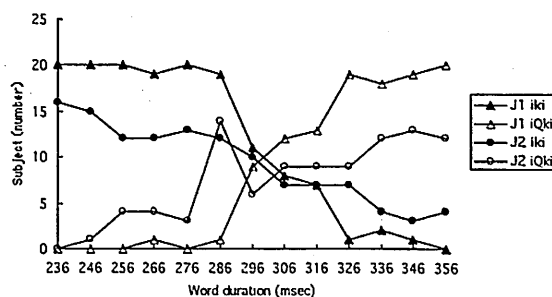
3. 実験結果

全体のラベリングはグラフ1-1からグラフ1-4のとおりである。グラフのY軸は当該トークンをラベリングした被験者の数、X軸はトークンの語全体の長さ (msec) である。凡例中 J1 は Yamane (1997) 中の、日本語ネイティブ・スピーカー20人の被験者のデータ、J2 は本論の日本語を第二言語として学習する被験者のデータである。

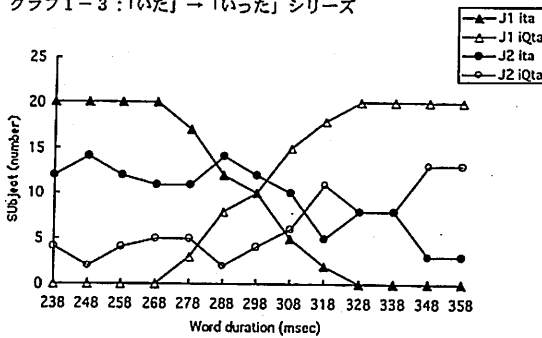
グラフ1-1: 「いき」→「いっき」シリーズ



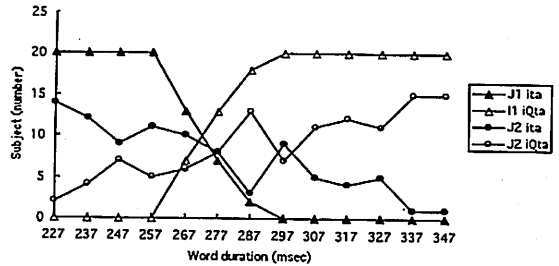
グラフ1-2: 「いっき」→「いき」シリーズ



グラフ1-3:「いた」→「いった」シリーズ



グラフ1-4:「いった」→「いた」シリーズ



各シリーズのグラフ中の線の交点、すなわち有促音と無促音の認識の境は、本論被験者 (J1) と Yamane (1997) の被験者 (J2) とで大差は見られない (詳細な数値は後述)。分布状態はこれら2グループでは異なっていて、J2のデータは規則性の点でよりノイズが見られる。実際、本論の実験では、日本語の通常の発話に存在する基本トークンに、そのターゲット・トークンと同値の語をラベリングする (例えば、「いき」

という基本トークンに「いっき」をラベリングする) 例が見られた。また、その逆に、ターゲット・トークンに基本トークンと同値の語をラベリングする例もあった。これらの件数は以下の表1のとおりである。表中「基→T」は上述前者の例、「T→基」は上述後者の例である。なお、本論被験者中少なくとも1回このようなラベリングを行ったのは16人中11人だった。

表1: ラベリングのノイズ

	iki → iQki		iQki → iki		ita → iQta		iQta → ita	
	基→T	T→基	基→T	T→基	基→T	T→基	基→T	T→基
J1 (N=20)	0	1	0	0	0	0	0	0
J2 (N=16)	6	2	0	3	4	3	1	2

これら2つのグループの結果を詳細に比較するために、境界値が定められた。境界値は、各々の被験者が、継続してターゲット・トークンと同値の語を認識し始めた中間トークンの数値から採られた。例えば、「いき」→「いっき」シリーズにおいて、ある被験者のラベリングが、トークン6以下全部=「いき」、トークン7=「いっき」、トークン8=「いき」、トークン9以上全部=「いっき」の場合は、トークン9が境界値とされた。各々の被験者の境界値の平均が、語全体の長さとし子音部の長さのそれぞれについてまとめられた。それぞれ順番に、表2-1、表2-2のとおりである。表中には、Yamane (1997) の実験で得られた、日本語のネイティブ・スピーカーによって通常のテンポで発話されたものの平均値もあわせて記されている。

J2のいずれの数値も、Yamane (1997) の発話実験で測定された無促音と有促音の値の間に存在している。つまり、本論被験者の促音認識上の境界値は、平均としては、日本語のネイティブ・スピーカーが自然

に発話する「いき」と「いっき」に見られるようなペアの値の間に、逸脱することなく存在しているということになる。さらに、どの数値においても、本論被験者と日本語ネイティブ・スピーカー (Yamane 1997) の間に統計的有意差は見られなかった^{iv)}。すなわち、日本語を第二言語として学習する成人は、語全体の長さとし子音部の長さに関しては、ネイティブ・スピーカーと差のない促音認識システムを有していると言える。

4. 結論と考察

仮説 (6) は否定された。本論の成人第二言語学習者のラベリングと、ネイティブ・スピーカーのラベリングとの間に有意差はなかった。このことは、音声タイミングの習得は、臨界期を過ぎても可能であるという論を支持している。

ただし、本論の実験結果は、これが全ての成人第二言語学習者にあてはまるということを必ずしも示唆し

表2-1: 通常発話における語全体の長さの平均値 (VCV 及び VQCV) と認識実験における境界値の語全体の長さの平均値 (J 1=日本語ネイティブ・スピーカー 及び J 2=日本語第二言語学習者) (msec) 表中 () 内の数値は標準偏差

	VCV (Yamane 1997)	J 1 (N=20)	J 2 (N=16)	VQCV (Yamane 1997)
iki → iQki	「いき」	265.0 (18.38)	302.7 (20.5)	「いっき」
iQki → iki	205.69 (23.94)	280.0 (19.03)	268.3 (20.9)	314.98 (33.11)
ita → iQta	「いた」	304.0 (14.29)	332.6 (19.0)	「いった」
iQta → ita	234.62 (12.81)	267.5 (9.44)	252.0 (21.4)	352.86 (25.70)

表2-2: 通常発話における子音部の長さの平均値 (C 及び QC) と認識実験における境界値の子音部の長さの平均値 (J 1=日本語ネイティブ・スピーカー 及び J 2=日本語第二言語学習者) (msec) 表中 () 内の数値は標準偏差

	C (Yamane 1997)	J 1 (N=20)	J 2 (N=16)	QC (Yamane 1997)
iki → iQki	/k/	116.0 (18.38)	153.7 (20.5)	/Qk/
iQki → iki	77.00 (12.74)	106.0 (19.03)	94.3 (20.9)	175.82 (24.35)
ita → iQta	/t/	175.0 (14.29)	203.6 (19.0)	/Qt/
iQta → ita	98.34 (11.76)	150.0 (9.44)	135.0 (21.4)	223.26 (24.49)

ない。表1のとおりノイズは少なく、ラベリングには個人差がある。これらの被験者の日本語が、未だに中間的言語 (interlanguage) の状態にあるのか、彼らが本実験では対象にしなかった何らかの別の音声の手がかり (例えば有促音の子音の前で強まる氣息 (aspirate) など) に依存しているのか、または音声でなく文脈に依存しているのかは本論の実験からは不明である。この成人第二言語学習者の個人差は、臨界期というものを越えたところで論を展開しようとする成人第二言語習得論 (adult second language acquisition) の最大の謎のひとつである。

参考文献

- Carr, P. (1999) English Phonetics and Phonology. Blackwell.
- Chomsky, N. (1965) Aspects of the Theory of Syntax. MIT Press.
- Chomsky, N. (1981) Lectures on Government and Binding. Foris.
- Curtiss, S. (1977) Genie. Academic Press.
- Eimas, P. (1975) Auditory and phonetic coding of the cues for speech: Discrimination of the [r-l] distinction by young infants. Perception and Psychophysics 18: 341-347.
- Han, M. (1994) Acoustic manifestations of mora timing in Japanese. Journal of the Acoustical Society of America 96: 73-82.
- Krashen, S. (1973) Lateralization, language learning and the critical period: Some new evidence. Language Learning 23: 63-74.
- Lenneberg, E. (1967) Biological Foundation of Language. Wiley.
- Miyawaki, K., Strange, W., Verbrugge, R., Jenkins, J. and Fujimura, O. (1975) An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. Perception and Psychophysics 18: 331-340.
- Pisoni, D., Lively, S. and Logan, J. (1994) Perceptual learning of nonnative speech contrast. In Goodman, J. and Nusbaum, H. (eds.), The Development of Speech Perception. MIT Press.
- Port, R., Dalby, J. and O'Dell, M. (1986) Evidence for mora timing in Japanese. Journal of the Acoustical Society of America 81: 1574-1585.
- Strange, W. and Dittman, S. (1984) Effects of

discrimination training on the perception of /r-l/ by Japanese adults learning English. Perception and Psychophysics 36: 131-145.

Yamane, M. (1997) Sounds of silence: The Japanese moraic consonant Q: Its status as a temporal cue in speech production and speech perception. ms. University of Connecticut.

- i 本論でいう「第二言語」とは、話者が言語活動の際優勢に用いる第一言語（一般に言う「母国語」）以外の言語を指す。ここでは、一般に言う「外国語」とほぼ同義に用いる。
- ii 本論でいう「言語」とは、人間の生得的な言語習得装置 (language acquisition device) で習得される内的言語 (internalized language) を指す

(Chomsky 1965, 1981 他)。ここには、言語習得装置とは独立した一般学習機能によって学習される規範文法、レトリック、綴り字に関する知識などは含まれない。

- iii (4) ではわかり易さを旨として、IPA表記でなくヘボン式ローマ字表記を用いた。なお、長音(「ー」で表記される長い音)は、便宜上ここでは母音扱いした。
- iv 「いき」→「いっき」シリーズ: $tob = 0.696 < tcv = 2.040$ ($df = 31, \alpha = 0.05$); 「いっき」→「いき」シリーズ: $tob = 0.210 < tcv = 2.040$ ($df = 31, \alpha = 0.05$); 「いた」→「いった」シリーズ: $tob = 0.626 < tcv = 2.040$ ($df = 31, \alpha = 0.05$); 「いった」→「いた」シリーズ: $tob = 0.350 < tcv = 2.038$ ($df = 32, \alpha = 0.05$).